DOCUMENT RESUME

ED 481 715 TM 035 094

AUTHOR	Moskal, Barbara M.
TITLE	Developing Classroom Performance Assessments and Scoring Rubrics - Part II. ERIC Digest.
INSTITUTION	ERIC Clearinghouse on Assessment and Evaluation, College Park, MD.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.; National Science Foundation, Arlington, VA.
REPORT NO	EDO-TM-03-03
PUB DATE	2003-06-00
NOTE	4p.; For part 1, see TM 035 093.
CONTRACT	ED-99-CO-0032;EEC-0230702
PUB TYPE	ERIC Publications (071) -- ERIC Digests in Full Text (073)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Criteria; *Performance Based Assessment; *Scoring Rubrics; *Test Construction; Test Interpretation; Test Results; Test Use
IDENTIFIERS	ERIC Digests

ABSTRACT

This Digest addresses developing scoring rubrics, administering performance assessments, and scoring, interpreting, and using results. In developing scoring rubrics, the criteria within the scoring rubric should be clearly aligned with the requirements of the task and the state goals and objectives. They should also be expressed in terms of observable behaviors or product characteristics. Scoring rubrics should be written in specific and clear language, with points that make sense, and a clear separation between score levels. It is also important that the statement of the criteria be fair and free from bias. Recommendations are also offered for administering performance assessments. Explanations of tasks should be clear and consider, and appropriate tools should be available to support the completion of the assessment activity. Scoring rubrics should be explained to students ahead of time. In scoring, interpreting, and using results, guidance is offered for using rates and anchor papers. The connection between the score or grade and the scoring rubric should be immediately apparent. Recommendations in this Digest can guide a teacher from planning, gathering, interpreting, and using, the four phases of classroom assessment. (Contains 10 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 481 715

Developing Classroom Performance Assessments and Scoring Rubrics - Part II

Barbara M. Moskal
Colorado School of Mines

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM035094

ERIC Digest EDO-TM-03-03 June 2003

BEST COPY AVAILABLE

Developing Classroom Performance Assessments and Scoring Rubrics - Part II

Barbara M. Moskal, Colorado School of Mines

A difficulty that is faced in the use of performance assessments is determining how the students' responses will be scored. Scoring rubrics provide one mechanism for scoring student responses to a variety of different types of performance assessments. This two-part *Digest* draws from the current literature and the author's experience to identify suggestions for developing performance assessments and their accompanying scoring rubrics.

This *Digest* addresses 1) Developing Scoring Rubrics, 2) Administering Performance Assessments and 3) Scoring, Interpreting and Using Results. Another *Digest* addresses Writing Goals and Objectives, and Developing Performance Assessments,

These categories guide the reader through the four phases of the classroom assessment process — planning, gathering, interpreting and using (Moskal, 2000a). The current article assumes that the reader has a basic knowledge of both performance assessments and scoring rubrics.

Developing Scoring Rubrics

Scoring rubrics are one method that may be used to evaluate students' responses to performance assessments. Two types of performance assessments are frequently discussed in the literature: analytic and holistic. Analytic scoring rubrics divide a performance into separate facets and each facet is evaluated using a separate scale. Holistic scoring rubrics use a single scale to evaluate the larger process. In holistic scoring rubrics, all of the facets that make-up the task are evaluated in combination. The recommendations that follow are appropriate to both analytic and holistic scoring rubrics.

Recommendations for developing scoring rubrics:

1. The criteria set forth within a scoring rubric should be clearly aligned with the requirements of the task and the stated goals and objectives. As was discussed earlier, a list can be compiled that describes how the elements of the task map into the goals and objectives. This list can be extended to include how the criteria that is set forth in the scoring rubric maps into both the elements of the task and the goals and objectives. Criteria that cannot be mapped directly back to both the task and the purpose should not be included in the scoring rubric.

2. The criteria set forth in scoring rubrics should be expressed in terms of observable behaviors or product characteristics. A teacher cannot evaluate an internal process unless this process is displayed in an external manner. For example, a teacher cannot look into students' heads and see their reasoning process. Instead, examining reasoning requires that the students explain their reasoning in written or oral form. The scoring criteria should be focused upon evaluating the written or oral display of the reasoning process.

3. Scoring rubrics should be written in specific and clear language that the students understand. One benefit of using scoring rubrics is that they provide students with clear description of what is expected before they complete the assessment activity. If the language employed in a scoring rubric is too complex for the given students, this benefit is lost. Students should be able to understand the scoring criteria.

4. The number of points that are used in the scoring rubric should make sense. The points that are assigned to either an analytic or holistic scoring rubric should clearly reflect the value of the activity. On an analytic scoring rubric, if different facets are weighted differently than other facets of the rubric, there should be a clear reason for these differences.

5. The separation between score levels should be clear. The scale used for a scoring rubric should reflect clear differences between the achievement levels. A scale that requires fine distinctions is likely to result in inconsistent scoring. A scoring rubric that has fewer categories and clear distinctions between these categories is preferable over a scoring rubric that has many categories and unclear distinctions between the categories.

6. The statement of the criteria should be fair and free from bias. As was the case with the statement of the performance activity, the phrasing used in the description of the performance criteria should be carefully constructed in a manner that eliminates gender and ethnic stereotypes. Additionally, the criteria should not give an unfair advantage to a particular subset of students that is unrelated to the purpose of the task.

Greater detail concerning the development of scoring rubrics, both analytic and holistic, is immediately available through this journal. Mertler (2001) and Moskal (2000b) have both described the differences between analytic and holistic scoring rubrics and how to develop each type of rubric. Books have also been written or compiled (e.g., Arter & McTighe, 2001; Boston, 2002) that provide detailed examinations of the rubric development process and the different types of scoring rubrics.

Administering Performance Assessments

Once a performance assessment and its accompanying scoring rubric are developed, it is time to administer the assessment to students. The recommendations that follow are specifically developed to guide the administration process.

Recommendations for administering performance assessments:

1. Both written and oral explanations of tasks should be clear and concise and presented in language that the students understand. If the task is presented in written form, then the reading level of the students should be given careful consideration. Students should be given the opportunity to ask clarification questions before completing the task.

2. Appropriate tools need to be available to support the completion of the assessment activity. Depending on the activity, students may need access to library resources, computer programs, laboratories, calculators, or other tools. Before the task is administered, the teacher should determine what tools will be needed and ensure that these tools are available during the task administration.

3. Scoring rubrics should be discussed with the students before they complete the assessment activity. This allows the students to adjust their efforts in a manner that maximizes their

BEST COPY AVAILABLE

performance. Teachers are often concerned that by giving the students the criteria in advance, all of the students will perform at the top level. In practice, this rarely (if ever) occurs.

The first two recommendations provided above are appropriate well beyond the use of performance assessments and scoring rubrics. These recommendations are consistent with the Standards of the American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999) with respect to assessment and evaluation. The final recommendation is consistent with prior articles that concern the development of scoring rubrics (Brualdi, 1998; Moskal & Leydens, 2000).

Scoring, Interpreting and Using Results

As was discussed earlier, a scoring rubric may be used to score student responses to performance assessments. This section provides recommendations for scoring, interpreting and using the results of performance assessments.

Recommendations for scoring, interpreting and using results of performance assessments:

1. Two independent raters should be able to acquire consistent scores using the categories described in the scoring rubric. If the categories of the scoring rubric are written clearly and concisely, then two raters should be able to score the same set of papers and acquire similar results.

2. A given rater should be able to acquire consistent scores across time using the scoring rubric. Knowledge of who a student is or the mood of a rater on a given day may impact the scoring process. Raters should frequently refer to the scoring rubric to ensure that they are not informally changing the criteria over time.

3. A set of anchor papers should be used to assist raters in the scoring process. Anchor papers are student papers that have been selected as examples of performances at the different levels of the scoring rubric. These papers provide a comparison set for raters as they score the student responses. Raters should frequently refer to these papers to ensure the consistency of scoring over time.

4. A set of anchor papers with students' names removed can be used to illustrate to both students and parents the different levels of the scoring rubric. Ambiguities within the rubric can often be clarified through the use of examples. Anchor papers with students names removed can be used to clarify to both students and parents the expectations set forth through the scoring rubric.

5. The connection between the score or grade and the scoring rubric should be immediately apparent. If an analytic rubric is used, then the report should contain the scores for each analytic level. If a summary score or grade is provided, than an explanation should be included as to how the summary score or grade was determined. Both students and parents should be able to understand how the final grade or score is linked to the scoring criteria.

6. The results of the performance assessment should be used to improve instruction and the assessment process. What did the teacher learn from the student responses? How can this be used to improve future classroom instruction? What did the teacher learn about the performance assessment or the scoring rubric? How can these instruments be improved for future instruction? The information that is acquired through classroom assessment should be actively used to improve future instruction and assessment.

The first three recommendations concern the important concept of "rater reliability" or the consistency between scores. Moskal and Leydens (2000) examine the concept of rater reliability in an article that was previously published in this journal. A more comprehensive source that addresses both validity and reliability of scoring rubrics is a book by Arter and McTighe (2001), Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance. The American Educational Research Association, American Psychological Association and National Council of Measurement in Education (1999) also address these issues in their Standards document. For information concerning methods for converting rubric scores to grades, see "Converting Rubric Scores to Letter Grades" (Northwest Regional Educational Laboratory, 2001).

Conclusions

The purpose of this article is to provide a set of recommendations for the development of performance assessments and scoring rubrics. These recommendations can be used to guide a teacher through the four phases of classroom assessment, planning, gathering, interpreting and using. Extensive literature is available on each phase of the assessment process and this article addresses only a small sample of that work. The reader is encouraged to use the previously cited work as a starting place to better understand the use of performance assessments and scoring rubrics in the classroom.

Acknowledgments

This article was originally developed as part of a National Science Foundation (NSF) grant (EEC 0230702), Engineering Our World. The opinions and ideas expressed in this article are that of the author and not of the NSF.

References

Boston, C. (Eds.). (2002). Understanding Scoring Rubrics. University of Maryland, MD: ERIC Clearinghouse on Assessment and Evaluation.

Brualdi, A. (1998). "Implementing performance assessment in the classroom." Practical Assessment, Research & Evaluation, 6(2) [On-line]. Available: http://ericae.net/pare/getvn.asp?v=6&n=2.

Mertler, C. A. (2001). "Designing scoring rubrics for your classroom." Practical Assessment, Research & Evaluation, 7(25). Available online: http://ericae.net/pare/getvn.asp?v=7&n=25.

Moskal, B. (2000a) "An Assessment Model for the Mathematics Classroom." Mathematics Teaching in the Middle School, 6 (3), 192-194.

Moskal, B. (2000b). "Scoring Rubrics: What, When and How?" Practical Assessment, Research & Evaluation, 7(3) [On-line]. Available: http://ericae.net/pare/getvn.asp?v=7&n=3.

Northwest Regional Educational Laboratory (2002). "Converting Rubric Scores to Letter Grades." In C. Boston's (Eds.), Understanding Scoring Rubrics (pp. 34-40). University of Maryland, MD: ERIC Clearinghouse on Assessment and Evaluation.

Perlman, C. (2002). "An Introduction to Performance Assessment Scoring Rubrics". In C. Boston's (Eds.), Understanding Scoring Rubrics (pp. 5-13). University of Maryland, MD: ERIC Clearinghouse on Assessment and Evaluation.

Rogers, G. & Sando, J. (1996). Stepping Ahead: An Assessment Plan Development Guide. Terra Haute, Indiana: Rose-Hulman Institute of Technology.

Wiggins, G. (1990). "The case for authentic assessment." Practical Assessment, Research & Evaluation, 2(2). Available online: http://ericae.net/pare/getvn.asp?v=2&n=2.

Wiggins, G. (1993). Assessing Student Performances. San Francisco: Jossey-Bass Publishers.

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract ED99CO0032. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education. **Permission is granted to copy and distribute this ERIC/AE Digest.**

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
X	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").